

Claim Amendments

1. (Currently Amended) A system for measuring an ability of a subject, comprising:

a first set of task items that require the subject to provide one or more spoken responses;

a speech recognition system coupled to receive the spoken response and to provide an estimate of the spoken response, the speech recognition system having an associated non-random inaccuracy;

a scoring device, the scoring device being operable to convert the estimate into [[an]] a task item score; and

a computation device, the computation device providing a subject score based on a combination of task item scores using a scoring computation model that depends upon an expected task item-dependent operating characteristic of the speech recognition system, including the associated non-random inaccuracy of the speech recognition system.

5

2. (Original) A system as claimed in claim 1, wherein the scoring computation model is based on Item Response Theory.

3. (Original) A system as claimed in claim 1, wherein the speech recognition system, the scoring device and the computation device comprise software modules running on a general purpose computing platform.

4. (Original) A system as claimed in claim 1, wherein the scoring computation model is constructed from a plurality of responses provided by a number of native and non-native speakers, the plurality of responses being prompted by a second set of task items.

5. (Original) A system as claimed in claim 1, wherein the estimate provided by the speech recognition system comprises an estimate of the linguistic content of the spoken response.

6. (Currently Amended) A system as claimed in claim 1, wherein at least one task item in the first set of task items tasks is ~~[[an]]~~ a task item selected from the group consisting of a prompt to read a sentence aloud, a prompt to repeat a word, a prompt to repeat a phrase, to repeat a sentence, a prompt to provide an opposite, and a prompt to answer a question.

7. (Currently Amended) In a computer-based system that grades spoken responses to a set of task items, wherein the system comprises a speech recognition system, an improved method of grading the spoken responses, the improvement comprising:

determining a subject score for the spoken responses to the set of task items, wherein the subject score accounts for ~~[[an]]~~ a task item-dependent operating characteristic of the speech recognition system to accurately recognize the spoken responses.

8. (Currently Amended) A method for measuring an ability of a subject, comprising:

providing a set of task items;

generating a difficulty value for each task item in the set, the difficulty value being based upon the task item and a performance measurement associated with an automatic device that measures task performance, wherein the performance measurement comprises a measure of ~~[[an]]~~ a task item-dependent operating characteristic of the automatic device from responses to the set of tasks;

obtaining a response to each task item from the subject; and

combining the difficulty values and the responses from the subject to form a subject score reflecting at least one of a linguistic ability and a cognitive ability of the subject.

9. (Canceled)

10. (Original) A method as claimed in claim 8, wherein the step of generating a difficulty value comprises the step of obtaining a plurality of sample responses from a group of sample speakers.

El 11. (Original) A method as claimed in claim 10, wherein the step of generating a difficulty value further comprises the step of applying a statistical model to the plurality of sample responses.

12. (Original) A method as claimed in claim 8, wherein the step of combining the difficulty values and the responses comprises the step of applying a statistical model to the plurality of responses.

13. (Canceled)

14. (Currently Amended) A method for measuring an ability of a subject comprising:
providing a set of task items tasks and a device that automatically measures performance of the task items tasks;

determining a difficulty value for each task item, wherein the difficulty value is based upon the task item and upon a performance measure associated with ~~[[an]]~~ a task item-dependent operating characteristic of an automated device in assessing performance of the task;

obtaining verbal responses to the task items ~~tasks~~ from the subject; and

combining the verbal responses and the difficulty values to form a subject score reflecting at least one of a linguistic ability and a cognitive ability of the subject.

15. (Original) A method as claimed in claim 14, wherein the device comprises an automated speech recognition system.

16. (Currently Amended) An apparatus for determining a difficulty value of task items in a test, comprising:

a set of responses to the task items from a number of individuals;

an automated grader, wherein the automatic grader ~~grades~~ receives the set of responses and provides graded responses; and

means for reducing the graded responses to a set of task item difficulties, said task item difficulties normalizing the task items by accounting for non-random errors by the automatic grader.

17. (Currently Amended) A method for determining a difficulty value of task items in a text, comprising:

obtaining a set of responses to the task items from a number of individuals;

automatically grading the set of responses, thereby generating graded responses; and

reducing the graded responses to a set of task item difficulties, said task item difficulties

including a measurement of [[an]] a task item-dependent operating characteristic associated with the

E) act of automatically grading the set of responses for purposes of normalizing the task items to provide an accurate assessment.
